MACRO-PANELS AND REALITY

Gianluca Cubadda  
Università di Roma "Tor Vergata"

Alain Hecq*  
University of Maastricht

Franz C. Palm  
University of Maastricht

April 3, 2007

Abstract

This note argues that large VAR models with common cyclical feature restrictions provide an attractive framework for parsimonious implied univariate final equations, justifying on the one hand the estimation of homogenous panels with dynamic heterogeneity and a common factor structure, and on the other hand the aggregation of time series. However, starting with a too restrictive DGP might preclude from looking at interesting empirical issues.

JEL: C32

Keywords: ARIMA, common cycles, panel data, final equations

*Corresponding author: Alain Hecq, University of Maastricht, Department of Quantitative Economics, P.O.Box 616, 6200 MD Maastricht, The Netherlands. Email: a.hecq@ke.unimaas.nl. Fax: +31 43 388.4874. Tel: +31 43 388.3798. Without any implications authors would like to thank Jean-Pierre Urbain for helpful comments.
1 A restrictive view of dynamic macro-panels

With the availability of many macroeconomic time series, one witnesses the application of panel data techniques to the modelling of macro dynamic relationships. For stationary as well as for non-stationary processes, pooling data for instance for \(N\) countries instead of considering one country often increases parameter efficiency and improves the power of some test statistics such as unit root tests. One restrictive representation of macro-panels inherited from time series is to consider a set of \(N\) AR(1) equations, we stack for notational convenience in the \(\text{VAR}(1)\)

\[
Z_t = \alpha + \Gamma_1 Z_t + \varepsilon_t, \quad t = 1..T, \quad (1)
\]

with \(Z_t = (Z_{1t},...,Z_{Nt})'\), \(\alpha = (\alpha_1,...,\alpha_N)'\) to account for individual fixed effects, \(\varepsilon_t\) has a covariance matrix \(\Sigma_e\) and \(\Gamma_1\) is an \(N \times N\) diagonal matrix with elements \(\phi_1,...\phi_N\). This framework accommodates numerous features such as a fixed individual effect, the possibility of some heterogeneity of the coefficients \(\phi_i\), or the contrary i.e. slope homogeneity if all \(\phi_i, i = 1...N\) are alike and some cross-sectional dependency if \(\Sigma_e\) is not diagonal.

In this paper we are interested in cases where a common coefficient \(\phi\) is obtained for all \(N\). In the presence of heterogeneity in the autoregressive parameters, a situation that makes sense for a panel of macroeconomic time series, Pesaran and Smith (1995) propose a mean group estimator (MG) such that \(\hat{\phi} = \frac{\sum_{i=1}^{N} \phi_i}{N}\). That statistics is based on the average of individual estimates and is consistent when \(T\) and \(N \to \infty\).

2 An alternative representation

We take the framework developed by Zellner and Palm (1974, 2004) and obtain the \(N\) univariate processes, called final equations derived from a possible large \(\text{VAR}(p)\)

\[
\Phi(L)Z_t = \alpha + \varepsilon_t, \quad (2)
\]

where \(\alpha\) is similar as before. Plemultiplying both sides of (2) by \(\Phi(L)^{adj}\), the adjoint of \(\Phi(L)\), yields

\[
\det[\Phi(L)]Z_t = \tilde{\alpha} + \Phi(L)^{adj} \varepsilon_t, \quad (3)
\]

with \(\tilde{\alpha} = \Phi(1)^{adj} \alpha\) a vector of constant terms in a fixed effect framework. From (3) each series is a finite order ARMA\((p^*,q^*)\), with the same lag structure and the same coefficients for the autoregressive part for every series, although the system is a finite order \(\text{VAR}(p)\).

However, a small order \(\text{VAR}\) with few series already generates univariate ARMA models with large \(p^*\) and \(q^*\), a finding that is rejected when tested on real data where one usually finds that quite parsimonious models with low autoregressive and moving average orders are appropriate. Indeed, an \(N\) dimensional \(\text{VAR}(p)\) would imply at most individual ARMA\((Np, (N-1)p)\) processes.\(^ 1\) However, Cubadda, Hecq and Palm (2007) show

\(^1\)This result generalizes to individual ARMA\((Np, (N-1)p + q)\) for VARMA\((p,q)\) processes.
that the existence of additional short-run reduced rank restrictions in the dynamics of the (possibly non-stationary) VAR($p$) delivers very parsimonious ARMA models. We call these restrictions common cyclical feature or cofeature restrictions. The strongest form of common cyclical features is the notion of serial correlation common feature (henceforth, SCCF) proposed by Engle and Kozicki (1993) and Vahid and Engle (1993). In this framework, the series $Z_{t}$ have $s$ SCCF relationships if there exists an $N \times s$ matrix $\delta$ with full column rank and such that $\delta \Phi(L)Z_{t} = \delta \alpha + \delta \varepsilon_{t}$ in (2). To illustrate the implication of SCCF for the implied univariate models we take a simple numerical example of a stationary VAR(1) with $N = 3$

$$Z_{t} = \begin{bmatrix} 0.5 & -0.5 & 0.5 \\ 0.25 & -0.25 + \omega & 0.25 \\ 0.5 & -0.5 & 0.5 + \omega \end{bmatrix} Z_{t-1} + \varepsilon_{t}. \quad (4)$$

In the general case with $\omega \neq 0$ the implied series of (4) are ARMA(3,2) because $\det[I - \Phi_{1}L] = -0.5L^{3}\omega^{2} + (1.25\omega + \omega^{2})L^{2} - (0.75 + 2\omega)L + 1$ and the adjoint has terms up to the lag polynomial 2. Now if $\omega = 0$, there exist $s = 2$ SCCF vectors in (4), so that (4) can be written in its factor representation

$$Z_{t} = \begin{bmatrix} 0.5 \\ 0.25 \\ 0.5 \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \end{bmatrix} Z_{t-1} + \varepsilon_{t}. \quad (5)$$

For (5) the final equations are ARMA(1,1) with the same autoregressive coefficient and a factor structure of rank $s$ on the MA part such that

$$(1 - 0.75L)Z_{t} = \varepsilon_{t} + \begin{bmatrix} -0.25 & -0.5 \\ 0.25 & -1 \\ 0.5 & -0.5 \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -0.5 \end{bmatrix} \varepsilon_{t-1}. \quad (6)$$

Generalizing that observation, Cubadda, Hecq and Palm (2007) prove that the maximal ARMA orders of the univariate series generated by a stationary VAR($p$) and a cointegrated VAR($p$) with $s$ SCCF relationships are respectively ARMA[(N - s)p, (N - s)p] and ARMA[(N - s)(p - 1) + r, (N - s)(p - 1) + r]. To give hints of these proofs, for instance, for the AR part in the stationary VAR($p$), let us rewrite equation (2) as $Q(L)X_{t} = \varepsilon_{t}$ where $X_{t} = MZ_{t}$, $\varepsilon_{t} = M\varepsilon_{t}$, $Q(L) = M\Phi(L)M^{-1}$, $M' \equiv [\delta, \delta_{\perp}]$, $M^{-1} = [\delta(\delta'\delta)^{-1}, \delta_{\perp}(\delta'_{\perp}\delta_{\perp})^{-1}]$ with $\delta_{\perp}$ is the orthogonal complement of $\delta$. It is clear that the maximum AR and MA orders of the univariate representation of elements of $Z_{t}$ must be the same as those of elements of $X_{t}$. It easily follows that $\det[Q(L)] = \det[\delta'_{\perp}\Phi(L)\delta_{\perp}(\delta'_{\perp}\delta_{\perp})^{-1}]$ is a polynomial of order $(N - s)p$.

Although convenient in terms of parsimony and economic interpretation of business cycles (cycles will be synchronous in the presence SCCF), assumptions underlying SCCF may be too strong. Indeed SCCF implies that all the coefficient matrices of the polynomial $\Phi(L) - I_{N}$ have a common left null space, a condition that has been relaxed inter alia in Ahn and Reinsel (1988), Tiao and Tsay (1989), Vahid and Engle (1997), in the weak form common feature model (WF, hereafter) by Hecq, Palm and Urbain (2006) or by Cubadda

3
and Hecq (2001). Detailed proofs on the orders for all these models as well as the form of the factor model for the VMA part in (6) are given in Cubadda, Hecq and Palm (2007) and are not the issue of this paper. The main points here is that a set of $N$ series with some common cyclical features may be represented as a parsimonious dynamic panel model that exhibits both an homogeneous AR part and cross-correlated VMA errors having a factor structure. Consequently, confounding the diagonal VAR (1) with the $N$ implied final equations of (3) or of the example in (6) is not without risk when the interest is in obtaining a common slope parameter. Next we shall consider several ways to obtain estimates of the AR part, both under homogeneity and under heterogeneity, and their implications for empirical research.

3 Mean group of individual estimation or aggregates?

We assume that $Z_t$ is stationary, $Z_t$ being possibly the first difference of underlying series having unit roots and generated either by (2) or by (1). Then, we identify and estimate by ML for each series individually the parsimonious empirical ARMA$(p_i,q_i)$ such that

$$Z_{it} = \hat{\alpha}_i + \sum_{j=1}^{p_i} \hat{\phi}_{ij} Z_{it-j} + \sum_{k=1}^{q_i} \hat{\theta}_{ik} \tilde{u}_{it-k} + \tilde{u}_{it}, \quad i = 1 \ldots N, \ t = 1 \ldots T,$$

(7)

where $\hat{\alpha}_i$, $\hat{\phi}_{ij}$ and $\hat{\theta}_{ik}$ are estimated scalar parameters for series $i$, $i \in \{1, 2, \ldots, N\}; p_i$ and $q_i$ are the lag orders of the ARMA for the $i$th series and they might empirically differ from series to series. In a second step one may compute the average of $N$ autoregressive parameters such that $\hat{\phi}_{ij}^{av} = N^{-1} \sum_{i=1}^{N} \hat{\phi}_{ij}$, $j = 1 \ldots \max(p_i)$.

Instead of estimating individual equations in (7) and because the implied model for $\det[\Phi(L)Z_t]$ is at most a MA[$(N-s)p]$ model in the presence of SCCF vectors for instance, we further use this observation and estimate an ARMA model for the average of the $N$ series in the ML estimation of

$$\bar{Z}_t = \bar{\alpha} + \sum_{j=1}^{p} \hat{\phi}_{j}^{av} \bar{Z}_{t-j} + \sum_{k=1}^{q} \hat{\theta}_{k} \bar{\tilde{u}}_{t-k} + \bar{\tilde{u}}_t,$$

(8)

with $\bar{Z}_t = N^{-1} \sum_{i=1}^{N} Z_{it}$ the simple average of $N$ series. This has the advantage to impose the common AR coefficients for all $N$ and to reduce the impact of the variability in the estimated parameters.

4 A Monte Carlo experiment

Let us now evaluate the small sample properties of these two strategies (7) and (8) on the two different types of DGPs. We only focus on the estimation of the common coefficients for the autoregressive part for $N$ fixed. The DGP 1 is the simple version (1) of DGPs typically used in the first generation of macro-panels, with an homogeneous autoregressive root such that $\Gamma_1 = diag(0.5)$. Introducing some heterogeneity does not alter our conclusions about the bias and only increases the RMSE. For $i = 1 \ldots N$, we have individual effects $\alpha_i$ which are chosen from a uniform distribution on $[0,1]$ and $\Sigma_c = I$. We consider successively $N = 5, 11, 23$ individuals for $T = 50, 100, 250$ data points. This corresponds to cross-sectional and time dimensions for instance for a set OECD or European countries.
Alternatively to the DGP 1, the DGP 2 is VAR(1) in (2) "without zeros" but with a reduced rank structure \((\text{rank}[\Phi_1] = 1)\) such that

\[
Z_t = \alpha + \Phi_1 Z_{t-1} + \varepsilon_t = \alpha + \delta_z C Z_{t-1} + \varepsilon_t.
\]

(9)

To keep fixed the AR and the MA degrees in the \(N\) implied models we impose that there exist \(s = N - 1\) SCCF relationships, leading to \(N\) implied ARMA(1,1) models whatever \(N\) instead of ARMA\((Np, (N-1)p)\). Our associated common feature matrix assumes full short-run convergence between variables (which could be GDP for different countries). Indeed, one of the issues we want to emphasize is to determine a core of countries (or variables) with the maximum number of co-movements. The cofeature matrix has thus the shape, for instance with \(N = 5\):

\[
\delta = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
-1 & -1 & -1 & -1
\end{pmatrix},
\]

leading to (up to a normalization) \(\delta_z = (1 : 1 : 1 : 1')\). \(\alpha\) is a \(N\) dimensional vector of constant terms that we generate as before. The factor \(C_1\) cannot change at each replication otherwise we could not compute a bias because implied roots would be different at each of these replications. In order to have the same AR coefficient whatever the odd number of individuals we choose \(C_0 = (0.5, -0.5, 0.5, -0.5, 0.5, ..., -0.5, 0.5)\).

This guarantees both the stationarity of the multivariate process and the common value of the AR root \((\phi_1 = 0.5, i = 1...N)\) for every series whatever the number of individuals. To show this we can simply use the property of partitioned matrices and compute \(\det[I - \Phi_1 z]\) when considering the \((N - 1) \times (N - 1)\) upper-left block of \((I - \Phi_1 z)\). For the covariance matrix of the VAR we use \(\Sigma = I\). This implies cross-correlated MA errors \(w_t = \Phi(L) \hat{\varepsilon}_t\) in the final equations because their contemporaneous covariance matrix is \(E(w_t w_t') = I + \Phi_1 \Phi_1'\). This corresponds to a contemporaneous correlation between the MA disturbances of the implied equations of \(\rho = 0.55\) for all pairs. We use \(M = 2000\) replications and generate \(T + 50\) observations by iteration before dropping the first 50 points to initialize the random sequence.

Different estimation strategies to computing a common autoregressive component are evaluated in terms of their empirical bias and root mean squared error which we define as

\[
\text{Bias} : \frac{1}{2000} \sum_{m=1}^{2000} (\phi_1^{(j)} - 0.5); \quad \text{RMSE} : \sqrt{\frac{1}{2000} \sum_{m=1}^{2000} (\phi_1^{(j)} - 0.5)^2},
\]

where \(\phi_1^{(j)}\) denotes one of the six estimator for the common AR(1) parameter of the \(N\) series. The first four procedures are based on the combination of individual estimations; estimators five and six are based on aggregates. In details these are:

\(\phi_1^{(1)}\) : \(N\) individual ARMA(1,1) models are estimated by ML before using the average \(\hat{\phi}_1^{(1)} = \frac{1}{N} \sum_{i=1}^{N} \hat{\phi}_{1i}\).
\( \hat{\phi}_1^{(2)} \): Here we average parameters from \( N \) AR(1) models, omitting the moving average part that can look non-significant in small samples. This estimator is obviously inconsistent for individual equations in the DGP 2 while it is the best strategy for DGP 1.

\( \hat{\phi}_1^{(3)} \): This third method is similar to the previous one but we estimate \( N \) AR(1) models by IV instead of OLS. For each equation we take as instruments their own lags from \( t-2 \) up to \( t-7 \).

\( \hat{\phi}_1^{(4)} \): This estimator is similar to the mean group in \( \hat{\phi}_1^{(2)} \) but to account for the effect of possible cross-sectional dependency we augment each autoregressive equation by \( \tilde{Z}_{t-1} \) (see Pesaran, 2006).

\( \hat{\phi}_1^{(5)} \): Instead of working on individual series we estimate an ARMA(1,1) model on the aggregates \( \tilde{Z}_t = \frac{1}{N} \sum_{i=1}^{N} \tilde{Z}_{it} \).

\( \hat{\phi}_1^{(6)} \): This estimator is similar to the previous one on aggregates but here an ARMA(1,1) or an AR(1) is preferred based on the minimization of SBC.\(^2\)

Each individual equation estimates \( \hat{\phi}_1^{(1)} \) and \( \hat{\phi}_1^{(3)} \) are consistent on \( T \) under both DGPs whereas the estimators \( \hat{\phi}_1^{(2)} \) and \( \hat{\phi}_1^{(4)} \) are consistent under DGP 1 but not in general under DGP 2. In particular \( \hat{\phi}_1^{(4)} \) assumes a particular factor structure which is not the one obtained by the final equations. \( \hat{\phi}_1^{(5)} \) and \( \hat{\phi}_1^{(6)} \) are consistent on \( T \) for \( N \) fixed and finite.

It emerges from Table 1, that the MG estimator based on the AR(1) model (\( \hat{\phi}_1^{(2)} \)) provides, as expected, the best results for DGP 1, i.e. the diagonal VAR. However, in the more general case of DGP 2, the best strategy is to estimate the parsimonious ARMA model (e.g. using SBC) on aggregates, i.e. estimates \( \hat{\phi}_1^{(6)} \). Other methods behave more poorly both in terms of empirical bias and accuracy (RMSE). In particular the augmented model performs very badly when applied to the final equations. The intuition behind all these results is that in the set of final equations we have an homogenous slope under the null and working with the averages allows to impose that restriction. This shows, if needed, that estimators to be used in panels heavily depend on the DGP. Augmenting the regression by the means of averages as suggested by Pesaran (2006) for a large \( T \) and large \( N \) panels gives an estimate of the diagonal of the VAR, which for the DGP 2 seems inappropriate.

5 Conclusion

Starting from an unrestricted multivariate VAR model we obtained the final equation representation which in the presence of serial correlation common features is very similar to models adopted in the panel literature, that is the models appear to have low order common AR parts, common factor representation resulting from the low degree MA parts and cross-sectional correlation. Starting from a diagonal as in the DGP in (1) or more generally block-diagonal VAR where the (block-)diagonality reflects the presence of intra-block/entity

\(^2\)In a VAR(\( p \)) with common features, and although the \( N \) implied processes are already parsimonious univariate ARMA(\( p*, q* \)), there exists a linear combination with \( 0 \leq \bar{q} < q* \). For instance, premultiplying (6) by \( (1 : -1 : 1) \), which is not the vector that would compute the average here, annihilates the MA component.
### Table 1: Estimation of a common autoregressive parameter

<table>
<thead>
<tr>
<th>DGP</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N/T</td>
<td>Bias</td>
<td>rmse</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>-0.131</td>
<td>-0.061</td>
</tr>
<tr>
<td>100</td>
<td>-0.050</td>
<td>-0.024</td>
</tr>
<tr>
<td>250</td>
<td>-0.080</td>
<td>-0.047</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>-0.060</td>
<td>-0.029</td>
</tr>
<tr>
<td>50</td>
<td>-0.136</td>
<td>-0.061</td>
</tr>
<tr>
<td>100</td>
<td>-0.079</td>
<td>-0.030</td>
</tr>
<tr>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>-0.126</td>
<td>-0.060</td>
</tr>
<tr>
<td>50</td>
<td>-0.049</td>
<td>-0.024</td>
</tr>
<tr>
<td>100</td>
<td>-0.077</td>
<td>-0.047</td>
</tr>
<tr>
<td>250</td>
<td>-0.058</td>
<td>-0.029</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>-0.125</td>
<td>-0.057</td>
</tr>
<tr>
<td>50</td>
<td>-0.075</td>
<td>-0.032</td>
</tr>
</tbody>
</table>

Note: The DGP 1 corresponds to the diagonal VAR(1) in (1); DGP 2 is a VAR(1) with cofeature restrictions. \( \gamma_1^{(1)}, \gamma_1^{(2)}, \gamma_1^{(3)} \) refer respectively to the MG parameters of N individual ARMA(1,1), AR(1)-OLS and AR(1)-IV. In \( \gamma_1^{(4)} \) the lagged average is added to account for a common factor before computing the MG. \( \gamma_1^{(5)} \) is from an ARMA(1,1) on the aggregates whereas in \( \gamma_1^{(6)} \) the parsimonious empirical ARMA is chosen on the aggregates using SBC.
dynamics but absence of inter-block/entity dynamics, there is no or little reason for assuming cross-entity parameter homogeneity and the presence of common factor structures requires an additional assumption.

In a time series framework the presence of co-movement in a large dimensional VAR has a natural economic interpretation and useful practical implications. Consequently it might be interesting to reconsider working on aggregates in order to detect turning points or to build business cycle coincident indicators for instance.

6 References


